

# AP Statistics- Worksheet 2.6D – Simpson's Paradox

Key

Suppose you need heart surgery and are trying to decide between two surgeons, Dr. Cardi and Dr. Kulp. Here are some data concerning their success rates:

	Survived	Died	Total
Dr. Kulp	800	200	1000
Dr. Cardi	900	100	1000
Total	1700	300	2000

- Calculate the proportion of Kulp's patients who survived and the proportion of Cardi's patients who survived. Which doctor saved the higher percentage of patients?

Kulp: 80%

Cardi: 90%

Suppose we further categorize each patient according to whether they were in good or poor condition prior to treatment. The results are as follows:

	Good Condition		
	Survived	Died	
Dr. Kulp	590	10	600
Dr. Cardi	870	30	900
	1460	40	1500

	Poor Condition		
	Survived	Died	
Dr. Kulp	210	190	400
Dr. Cardi	30	70	100
	240	260	500

- Verify that these two tables add together to form the table in the first question above.
- Among those who were in good condition, compare the recovery rates for the two doctors. Which doctor saved the greater percentage of its patients who had been in good condition?

Kulp: 98.3%

Cardi: 96.67%

- Among those who were in poor condition, compare the recovery rates for the two doctors. Which doctor saved the greater percentage of its patients who had been in poor condition?

Kulp: 52.5%

Cardi: 30%

- Write a few sentences explaining how Dr. Cardi had the higher recovery rate overall, yet Kulp has the higher recovery rate for each type of patient. (hint: did one doctor tend to treat more of one type of patient?)

- Cardi treated a lot ~~of~~ more patients in good condition

- These patients are easier to save

- Which doctor would you rather see if you were ill? Why?

Kulp!

This illustrates SIMPSON'S PARADOX.

This illustrates SIMPSON'S PARADOX.

- See other example p. 199 #2.32
- A conclusion is made when looking @ 2 variables; however the condition is reversed when we add a 3<sup>rd</sup> variable + break the data up more
- From wksh:
  - Cardi looked better overall
  - Kulp looked better (opposite concl. when we added the 3<sup>rd</sup> variable of "patient condition.")